EAI Endorsed Transactions on Game Based Learning

Editorial **ICST.ORG**

Quality Criteria for Educational Games

P.M. Escudeiro[1],*, N.F. Escudeiro[1], R.M. Reis[1], A. Barata[1] and R.Vieira[1]

[1]ISEP – GILT, Rua Dr. António Bernardino de Almeida, 431, 4200-072 Porto Portugal

Abstract

The development of software products in general and educational games in particular is a very demanding process involving costly resources for long periods of time. Evaluating its quality throughout the development lifecycle contributes to the early identification of imperfections and to improved efficiency.
In this paper we discuss a set of quality criteria and a method to assess the quality of educational games. These quality criteria are based on the functional requirements of the system and structured into two groups: general and specific. General criteria refer to the set of criteria that are common to all educational games in a given study area. Specific criteria refer to those criteria particularly addressing a specific educational purpose. We also present a quality model that is based on the software engineering paradigms. This model, called Quantitative Evaluation Framework, provides a quantitative measure for the quality of an educational game.

Keywords: Game-Based Learning**,** Assessment; Quantitative Evaluation; Educational Software, Educational Games, Quality Criteria Classification for games.

Received on 30 October 2012; accepted on 09 January 2013; published on 20 March 2013

doi: 10.4108/trans.gbl.01-06.2013.e3

1. Introduction

The development of digital learning content, especially the development of educational games, is a very demanding process requiring efficient control throughout its development cycle.
Our proposal allows assessing quantitatively the quality of a software product, in general, at a final pre-release stage as well as at any moment during the development lifecycle (Bates, 2000). In this paper we discuss the application of our proposal to a particular type of software product: educational games.
Our proposal is based on a quality measure addressing the user needs related to the educational game objective. The quality is based on a set of requirements that are grouped in general and specific requirements. These requirements constitute the set of criteria that will be used to assess the quality of the educational game (Scalet et al., 2000).
General requirements are those that support the functionalities that are common to any educational game. Specific requirements are those that assist the specific functionalities of each particular type of game being developed, such as adventure, action or serious games.
All the functional requirements of the educational game being developed are expected to be represented either as general or specific, thus contributing to assess the quality of the educational game. These are supposed to be identified and/or validated by an expert in the game design area at an early stage of the development.
Once identified, the quality criteria/requirements are assigned to factors. Each factor groups the subset of all the criteria that are relevant to it. A given requirement may be assigned to more than one factor. The relevance of a requirement to a concrete factor is established by a weight.
Factors on their own are grouped in dimensions. These dimensions define the quality space where the educational game will be evaluated.
A quantitative quality model is used to assess the quality of the educational game in this quality space at any stage during the development lifecycle. This model, called Quantitative Evaluation Framework (QEF), is suited to be applied to the development of any product. The quality space in QEF is adapted to a specific system though the

*Corresponding author. Email: pmo@isep.ipp.pt

identification of proper quality dimensions, factors and requirements (quality criteria).
In the remainder of this paper we describe the QEF methodology, in Sections 2 and 3, followed by a description of the quality criteria defined to assess educational games, in Section 4. In Section 5, we present the application of QEF to a specific educational game. In Section 6, we present our conclusions.

2. Evaluation model philosophy

A simple question for any educational software should be, ”can this product actually teach what it is supposed to?” It is a simple question to ask, but often difficult to answer because the product may have so many beguiling features. It requires the evaluator to recognize his/her own view of the way in which students learn, to relate that view to the learning objectives and to determine how and whether those objectives are accomplished by the educational game under development.
The application of QEF throughout the development lifecycle of a game highlights the flaws that are present in the current version at the time of evaluation (Bloom, 1964).. It allows the development team to focus on those flaws, guiding the development process towards the desirable outcome. Moderating the development of an educational game with QEF promotes the quality of the final product (Oliver, 2000).Moreover, it allows the early identification of flaws that can be corrected early in the development process thus avoiding unnecessary costs due to the effects of error propagation.
The QEF quality space is an n-dimensional Euclidean space. Each dimension represents one particular major quality aspect to consider taking into consideration the type of product being developed, its end use and objectives. For educational games we consider a three-dimensional space. The dimensions assumed to be of relevance for educational games are: Pedagogical; Technical and Organizational.
Every dimension aggregates a set of factors deemed to be relevant from its own perspective. A factor is a component that represents the system performance from a particular point of view. The dimensions and the factors that constitute the quality space are defined by the content expert, in our example by the Game Learning expert.
Quality is directly dependent on the perceived differences between deployed and designed functionality.

3. Quality assessment methodology

The quality of the educational software, especially educational games, depends on the context and the target aimed by the learning system. These characteristics set the relevant evaluation factors. In concrete, the quality of a given educational software system, q, is defined in a tri-dimensional quality space, Q. Quality is computed in comparison to an hypothetic ideal system, I, whose quality is assumed to be 100%.
The fitness of the deployed educational software system, with regards to its final purpose, is evaluated over a set of core properties that it must exhibit to fulfil its purpose and achieve its aims. These desirable properties are the evaluation criteria that we rely on.
The quality of a given system is computed from the distance between the ideal system, I, assumed to have the coordinates (1,1,1) in our quality space, and the coordinates representing the system under evaluation at the evaluation moment.

Each dimension aggregates a set of factors. The coordinates of the real system are weighted means of the factors contributing to each quality dimension, Dim_i.

$$Dim_i = \sum_n (p_n \times factor_n), \sum_n p_n = 1. \forall_n\, p_n \in [0,1] \qquad (1)$$

Where n is the number of relevant factors for the Dim_i dimension.
Each factor is computed by Equation 2 as the weighted mean of the fulfilment of the evaluation criteria directly contributing to it, pc_m. The weight of a given criterion, pr_m, is its relevance to the factor. These weights, defining the ideal system, are assigned by the content expert during the design stage of the educational game.

$$Factor_n = \frac{1}{\sum_m pr_m} \times \sum_m (pr_m \times pc_m) \qquad (2)$$

The global deviation of the real system, which is represented by the coordinates (Dim_1, Dim_2, Dim_3) in our quality space, w.r.t. the ideal system, represented by the coordinates (1,1,1), is computed by the Euclidean distance between these two points in the quality space, according to Equation 3.

$$D = \sqrt{\sum_j \left(1 - \frac{Dim_j}{100}\right)^2} \qquad (3)$$

The real system quality is then computed by Equation 4:

$$Q = 1 - \frac{D}{\sqrt{n}}, Q \in [0,1] \qquad (4)$$

Or, in percentage (Equation 5):

$$q = \left(1 - \frac{D}{\sqrt{n}}\right) \times 100, q \in [0,100] \qquad (5)$$

When defined this way, the quality, Q, is in the inverse proportion of the distance, D, between the ideal system and the real system, measured in the quality space. When D=0, then Q=1, and when $D=D_{max}=\sqrt{n}$, then Q=0.

3.1. Quality measurement process

The process to measure the quality of a given system includes the following steps:

- Setup phase, under the responsibility of the content experts
    1. Definition of the relevant dimensions of the quality space.
    2. Definition of the relevant factors for each dimension.
    3. Definition of the relevant requirements/ evaluation criteria given the system context and aims.

4. Tuning of weights, including p_n, the weight of factors to dimensions and pr_m, the weight of requirements/evaluation criteria w.r.t. factors.

- Deployment phase, under the responsibility of the content experts, system evaluators and development team

1. Evaluation of the fulfillment percentage of each requirements/evaluation criteria, pc_m.
2. Compute the performance achieved by each dimension (Equation 1).
3. Compute global deviation in quality space (Equation 3).
4. Compute quality (Equations 4 and 5).

Setup phase

In this phase, previous to any evaluation, we must define the ideal system, our golden standard for the system/content to be evaluated.

Dimensions and Factors

The definition of the quality space structure, including the quality dimensions and factors, are determined by the field domain and context of the system/content being developed. Once defined, the same quality space structure can be used to evaluate any system in the same domain.

Relevance of factors to dimensions

For each dimension we must set the importance of the relevant factors. Each factor being relevant to a dimension will contribute to it with a weight that is proportional to its importance for the dimension. These weights are such that each has a value between 0 (the factor is not relevant to the dimension) and 1 (the factor is the only one being relevant to the dimension). The weights associating factors to a dimension sum to 1 (see Equation 1 above).

Factors and evaluation criteria

The requirements/evaluation criteria might be relevant or not to each factor. Their relevance are more dependent on the specific system/content being developed than factors or dimensions. There are probably some generic criteria, that are valid for all the systems in a given domain, but there are probably also many more that are specific to a given system in particular. These requirements/evaluation criteria must be set by experts in the domain and in the specific area addressed by the system under evaluation.

The criteria weights w.r.t. factors are volatile with a rather subjective nature. To assure some robustness to the quality measure we have restricted these weights to a limited set of discrete values ranging between 0 (non-relevant criteria) and 10 (core criteria) according to the following scale:

10- Fundamental, crucial
8 – Very important
6 – Important
4 – Necessary
2 – Optional
0 – Non-relevant

3.2. Deployment phase

Having the evaluation framework previously set at the setup phase, including the quality space structure (dimensions and factors) and the corresponding evaluation criteria, we may now use it to evaluate the system and compute its quantitative quality value.

The evaluation is based on the percentage of fulfilment that is achieved by the system under evaluation for each of the evaluation criteria previously set.

The task to be performed by the evaluators of the system is the assignment of these fulfilment percentages to each of the evaluation criteria. Once this is performed, the rest is automatic and the several quality indicators become be immediately available.

It is very important to have a very good selection of the evaluation criteria. These criteria, when evaluated, should provide a correct view of the system performance.

4. Quality criteria for games

Regarding to educational software in general we have had previously establish a common definition for the quality space (Escudeiro and Bidarra, 2006)that is adapted in this paper for educational games in particular.

The dimensions of our quality space for educational games in particular are: Pedagogical; Technical and Organizational.

The **Pedagogical dimension** reflects the characteristics of the educational game related to the teaching/learning procedure supported by the learning criteria. Learning is determined by several factors involving the interrelationship between individuals. This dimension aggregates three factors: Pedagogical Approach; Content Organization and Trainee's tools.

The **Technical dimension** reflects the technical characteristics of the educational game related to software design and development. These are the characteristics that might guarantee overall performance, easy and safe use (Wisner, Fialho & Santos 1995) i.e., the aspects that are directly related to the technical and ergonomic aspects of the educational game. This dimension adds, among others, the following factors: User Interface; Interaction; Game Play and Socio-Cultural aspects.

The **Organizational dimension** reflects the characteristics of the educational game regarding to its operational aspects.

This dimension is determined by one factor: Administration Tools.

Table 1 presents the general quality criteria that contribute to each evaluation factor. General quality criteria are those supporting common functionalities to any educational game.

Table 1. General quality criteria for educational games

Dimensions	Factor	Quality Criteria
Pedagogical	Pedagogical Approach	Learning objectives are clearly defined and focused on the specific learning domain to be addressed (eventually referring to Bloom's taxonomy: Cognitive, Affective, and Psychomotor)

Pedagogical	Pedagogical Approach	Self- assessment tools are available (diagnostic evaluation, formative evaluation, summative evaluation and self-assessment)
		Self-learning training is provided if required
		The pedagogical prerequisites were defined
		The game can be integrated in different pedagogical methodologies
	Content Organization	The content is valid and trustworthy.
		The content corresponds to the target group needs
		The content is effective addressing the learning objectives.
		The presentation of concepts is clear and timely
		The organization, structuring and sequencing of content supports an effective learning process
		Written/spoken content is free of grammatical and syntactical errors.
		The text is clear, simple and legible
		The technical terminology of the field being taught is properly applied
		The symbols and/or metaphor is appropriate and consistent
	Trainee's tools	Trainees are allowed to use online help
		Each trainee has access to his/hers grading and progress
		Each trainee has access to his/hers grading and progress

Technical	User Interface	The items presented in the user interface (screens) contribute to motivate the trainees
		Text enhances information
		Graphics and pictures provide useful information
		The sound used enhances information
		The video used enhances information
		The communication speed between user and game is adequate
	Interaction	The interaction with the system is intuitive for the target audience
		The definition of technical expertise required from the user is provided
		The instructions are clear, accurate, and concise
		The user can play the game without reading the manuals
		The player can change the difficulty level
		Feedback of user actions is provided
	game play	Navigation is consistent throughout the course and is carried out easily
		Multiple types of navigation are supported (e.g. for beginners, advanced, etc.)
		The basic navigation activators (buttons) are included in each screen (e.g. next, previous page, return to home)
		A help button is provided
		The game has no runtime errors

Technical		The software has an efficient programmatic structure
	Socio-Cultural aspects	The game is easily converted to several languages
		No offensive content or stereotypes in terms of gender, race, religion, cultural diversity is included
		No content inciting to violence is included
		The language used is suitable for various races and cultures
Organizational	Administration tools	Crash recovery tools are included
		Backup procedures are included
		There are tools that prohibit unauthorized access to the game
		The administrator is able to determine access rights for trainers and trainees
		Usage statistics are provided
		Game content can be efficiently updated
		The number of simultaneous online users is configurable
		Terms and conditions for using the learning resources are included
		Remote access is included

These generic requirements together with the specific requirements still under validation are being the base for the design of the educational game being presented in this paper.

5. Game engine supporting the game-learning document

The general quality criteria presented above are defined and validated during the initial stage of the development lifecycle of the educational game. In parallel, the development team produces the design document, which presents the analysis and design of the educational game to be developed.

The game engine supporting the final product is not a constraint for the application of QEF. In our case, the game is being developed using the e-Adventure game engine (Adventure Game Engine - Copyright © 2012 Sylvain Seccia) This game engine is directed to assist the development of adventure educational games.

The development lifecycle (Figure 1) is guided by a set of learning objectives previously established. These objectives supply guidelines to define the programming classes used in the development of the educational game (Figueiredo 2005).

The first stage of the development lifecycle is to set the target audience and then the general learning objectives. In the following step the specific learning objectives are set on the basis of the previously defined general objectives.

Once the specific learning objectives are set, we start to build the Game-Learning Document. This document establishes a relationship between the Pedagogic Domain and the Game Domain.

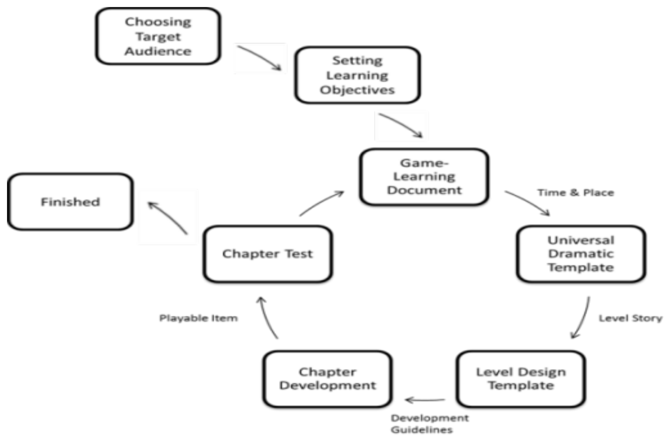

Figure 1. Development lifecycle

The Game-Learning Document is built according to the model supported by the e-Adventure game engine. By filling in this document the developers assure a direct association between the learning objectives and the technical development supported by this game engine. Figure 2 shows exactly the breakdown of these elements and how they are associated.

Figure 2. Relationship between pedagogic and game domains

The model also contemplates the hidden curriculum associated with gaming and pedagogical activities. Its objectives are deeply associated not with the game structure in itself but rather with graphics, audio and interaction Figure 3.

Annex I shows a sample of a Game-Based Learning Document, and of what can be extracted from it.

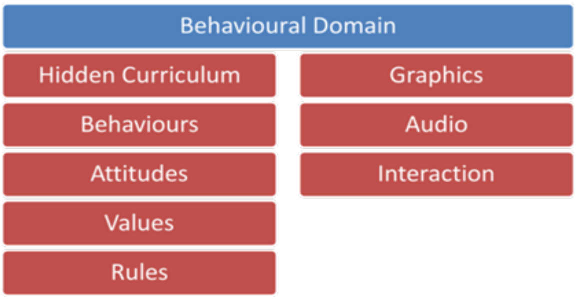

Figure 3. Behavioural domain relationship with digital game elements

6. Conclusion

This paper discusses a model and a set of criteria aimed to assess the quality of educational games. These criteria were established to assist the quality control of any educational game throughout its development lifecycle. Quality criteria are organized in two distinct groups: general and specific requirements.

These criteria are the base of a quality model that supports a quantitative assessment of the quality of any educational software. We show how to tailor it to the specific case of educational games. We demonstrate the application of this model to control the quality of a specific educational game designed to stress the importance of digital arts to present concepts like architecture in the Old City of Jerusalem.

Acknowledgements.

This research was supported by GILT (Graphic Interaction and Learning Technology).

References

[1] A. Dias Figueiredo (2005), Context Engineering: And is Development Research Agenda, Universidade Coimbra, 2005

[2] Bates Tony, (2000), A. W. Tony Managing Technological Change: Strategies for College and University Leaders. San Francisco, 2000

[3] Bloom Bertram B. Mesia, and David R. Krathwohl, (1964), Bloom B. S. Taxonomy of Educational Objectives,: The Affective Domain & The Cognitive Domain. New York: David McKay Co Inc.

[4] Escudeiro, Paula; Bidarra José, (2006), X-TEC: Techno Didactical Extension for Instruction/Learning Based on Computer – A new development model for educational software, Setúbal, Portugal, WEBIST 2006.

[5] Oliver 2000, Oliver M, Evaluating online teaching and learning, Information Services and Use, vol 20 (2/3)

[6] , Prentice Hall, Englewood Cliffs, New Jersey, 1991.

[7] Scalet et al, 2000, ISO/IEC 9126 and 14598 integration aspects: The Second World Congress on Software Quality, Yokohama, Japan, 2000.

[8] W, Fialho & apoud Santos 1995 Manual de análise ergonômica no trabalho. Curitiba:Gênesis, 1995.

ANNEX I - Game-Based Learning Document

<table>
<tr><td colspan="5">GAME-BASED LEARNING DOCUMENT</td></tr>
<tr><td colspan="2">Pedagogical Domain</td><td colspan="3">Game Domain</td></tr>
<tr><td>Didatical Unit</td><td>Digital Art</td><td>Game</td><td colspan="2">Serious Game about Digital Art</td></tr>
<tr><td colspan="2">General Learning Objective</td><td>Know the sub-part of the digital art process development process concerning its exhibition placement</td><td>Chapter Name</td><td>Find the spot</td></tr>
<tr><td colspan="2" rowspan="3">Specific Learning Objective</td><td>Locate a suitable place for the exhibition of a digital art artifact knowing its essence and associated restrictions</td><td rowspan="3">Action/Item</td><td>Answer the phone call to know the essence and restrictions associated with the artifact exhibition</td></tr>
<tr><td rowspan="2">Take the place location to the producers of the artifact</td><td>Gather the GPS coordinates of the exhibition place</td></tr>
<tr><td>Go to the production studio and deliver the coordinates</td></tr>
<tr><td colspan="5">Hidden Curriculum</td></tr>
<tr><td colspan="4">Recognize the typical architecture of the Old City of Jerusalem</td><td>Background based on pictures of the Old City of Jerusalem</td></tr>
<tr><td colspan="4">Quotidian movement of the Old City of Jerusalem</td><td>Typical sounds and images of people and places of the Old City of Jerusalem</td></tr>
</table>